

## SUPPORTING ONLINE MATERIAL

### Materials and Methods

**CAN1 forward mutation assays.** All data (with the exception of the WT spectrum from Lippert *et al.* 3) were derived using isogenic derivatives of strain  $\Delta|(-2)|\text{-}7\text{B}\text{-YUNI300}$  (*MATa CAN1 his7-2 leu2-}\Delta\text{::kanMX ura3-}\Delta\text{ trp1-289 ade2-1 lys2-}\Delta\text{GG2899-2900; ref. 18). The *rnh201* $\Delta$  strain was described previously (1). The *TOP1* ORF was deleted via transformation with a PCR product amplified from pAG25 (19). Cells were grown nonselectively at 30°C in YEP medium (1% yeast extract, 2% bacto-peptone; 2% agar for plates) supplemented with 2% dextrose and 250 µg/ml adenine. Can-R mutants were selected on synthetic complete medium lacking arginine and supplemented with 60 µg/ml canavanine. Rates of spontaneous Can-R colonies were measured as described previously (20). To generate spectra, genomic DNA isolated from independent Can-R mutants was PCR-amplified and sequenced. Rates of individual mutation types were calculated by multiplying the proportion of the mutation type in the corresponding spectrum by the total mutation rate.*

**lys2 reversion assays.** All *lys2* alleles were located at the *HIS4* locus of strain YPH45 (*MATa ura3-52 ade2-101<sub>oc</sub> trp1Δ1*; ref. 21) and were transcribed in the same direction as the replication fork initiated at *ARS306*. The *lys2ΔBgI* mutation (9) was introduced into the *his4Δ::pTET-LYS2* allele (22) by two-step allele replacement. All *lys2ΔA746NR* alleles were under control of the native *LYS2* promoter, and construction of the *lys2ΔA746NR,(AT)<sub>2</sub>* and *lys2ΔA746NR,(TC)<sub>3</sub>* alleles was described previously (3). The *lys2ΔA746NR,(AG)<sub>4</sub>* allele was similarly constructed using oligos 5'-GATCAAGTGAAGAGAGAGCTTAcGCAA and 5'-GATCTTGCgTAAGCTCTCTTCACTT; runs are underlined and lowercase letters reflect an AT to CG mutation introduced to remove a stop codon. The *RNH201* and *TOP1* genes were

deleted using PCR-amplified fragments containing dominant drug-resistant or nutritional markers as appropriate. Lys<sup>+</sup> rates and reversion spectra were obtained as described previously (22). For nonselective growth of *pTET-lys2ΔBgl* strains, medium was supplemented with 2 µg/ml doxycycline hyclate (Sigma) to maintain a low-transcription status.

**Top1 cleavage assays.** Cleavage assays were carried out as described previously (23). Briefly, custom-synthesized oligonucleotides (Midland Certified, Midland, Texas) were end-labeled prior to annealing with the cold, complementary strand at a 1:1 ratio. Reaction mixtures contained 50-200 nM of labeled DNA construct, 10 mM Tris-HCl (pH=7.5), 50 mM KCl, 1 mM EDTA, 1 mM DTT, 15 ug/mL BSA and 10% DMSO. Recombinant human Top1 (24) and/or CPT were added to 70 nM and 10 µM, respectively, as indicated. Reactions were for 1 hr at 25°C and were stopped by addition of 0.5% SDS (final concentration). In reversal assays, samples were incubated for 1 hr at 25°C before adding 0.5 M NaCl (final concentration). A small volume was withdrawn from the sample and the reaction stopped by adding 0.5% SDS (final concentration) at the specified time points (in min). Samples were analyzed on 20% denaturing polyacrylamide gels and labeled fragments detected using a Phospholmager.

**Table S1. Mutation types and rates in *lys2* frameshift-reversion assays**

<i>lys2</i> frameshift	Relevant allele	genotype	Rate (x 10 <sup>-9</sup> )				
			Hotspot	1-bp indels	Other	Total (CI)	Total/WT
$\Delta A746NR$	WT*	(N=90)	NA	0.14	0.81	0.95 (0.86-1.04)	
$\Delta A746NR, (AG)_4$	WT (N=84)		0.65	0.47	1.92	3.05 (2.58-4.22)	1.0
	<i>rnh201Δ</i> (N=95)		158	<2.0	9.9	189 (178-200)	62
	<i>rnh201Δ top1Δ</i> (N=88)		0.16	0.59	2.11	2.86 (1.68-3.61)	0.9
$\Delta A746NR, (TC)_3$	WT* (N=74)		1.43	0.19	1.17	2.79 (2.46-3.10)	1.0
	<i>rnh201Δ</i> (N=92)		17.7	<0.19	1.92	19.9 (15.7-22.7)	7.1
	<i>rnh201Δ top1Δ</i> (N=74)		<0.02	0.41	1.12	1.53 (1.34-1.89)	0.5
$\Delta A746NR, (AT)_2$	WT* (N=80)		6.21	0.56	0.65	7.42 (6.96-7.57)	1.0
	<i>rnh201Δ</i> (N=77)		8.15	0.44	2.62	11.2 (10.3-11.9)	1.5
	<i>rnh201Δ top1Δ</i> (N=74)		<0.03	0.73	1.21	1.94 (1.38-5.06)	0.3
$\Delta BglI$	WT (N=73)		0.22	1.66	1.31	3.19 (2.63-4.73)	1.0
	<i>rnh201Δ</i> (N=84)		5.65	0.57	0.57	6.88 (4.43-12.0)	2.2
	<i>rnh201Δ top1Δ</i> (N=91)		<0.05	1.05	3.74	4.80 (4.09-6.14)	1.5

Table legend

Rates of individual mutation types were calculated by multiplying the total reversion rate by the proportion of the mutation type in the corresponding spectrum. When no events of a given type were observed, the rate is designated as “<” and was calculated assuming a single event.

NA, not applicable; CI, 95% confidence interval.

\*Data are from Lippert *et al.* (3).

### Supplemental Figure legends

**Figure S1.** Mutation spectra of Can-R mutants. Nucleotides are numbered beginning with the ATG start codon. Base substitutions and indels are in red below and above the sequence, respectively; the lengths of the red bars correspond to deletion sizes.

**Figure S2. Reversion spectra for *lys2* alleles.** In **A-C** hotspot-containing sequences transplanted from the *CAN1* locus are highlighted in yellow and sequences associated with the flanking *Bgl*II site are highlighted in gray. + or Δ correspond to the insertion or deletion, respectively, of a single bp; cins or cdel, insertions or deletions associated with one or more base substitutions; del, deletion; dup, duplication; N, number of independent revertants sequenced.

**Figure S3. Human Top1 cleavage assays.** **A** shows fragments used as substrates, with relevant dinucleotide repeats highlighted in gray. Transcribed and nontranscribed strands are designated TS and NTS, respectively; arrows indicate positions of Top1 cleavage. In **B**, the indicated strand was labeled on the 3' end.

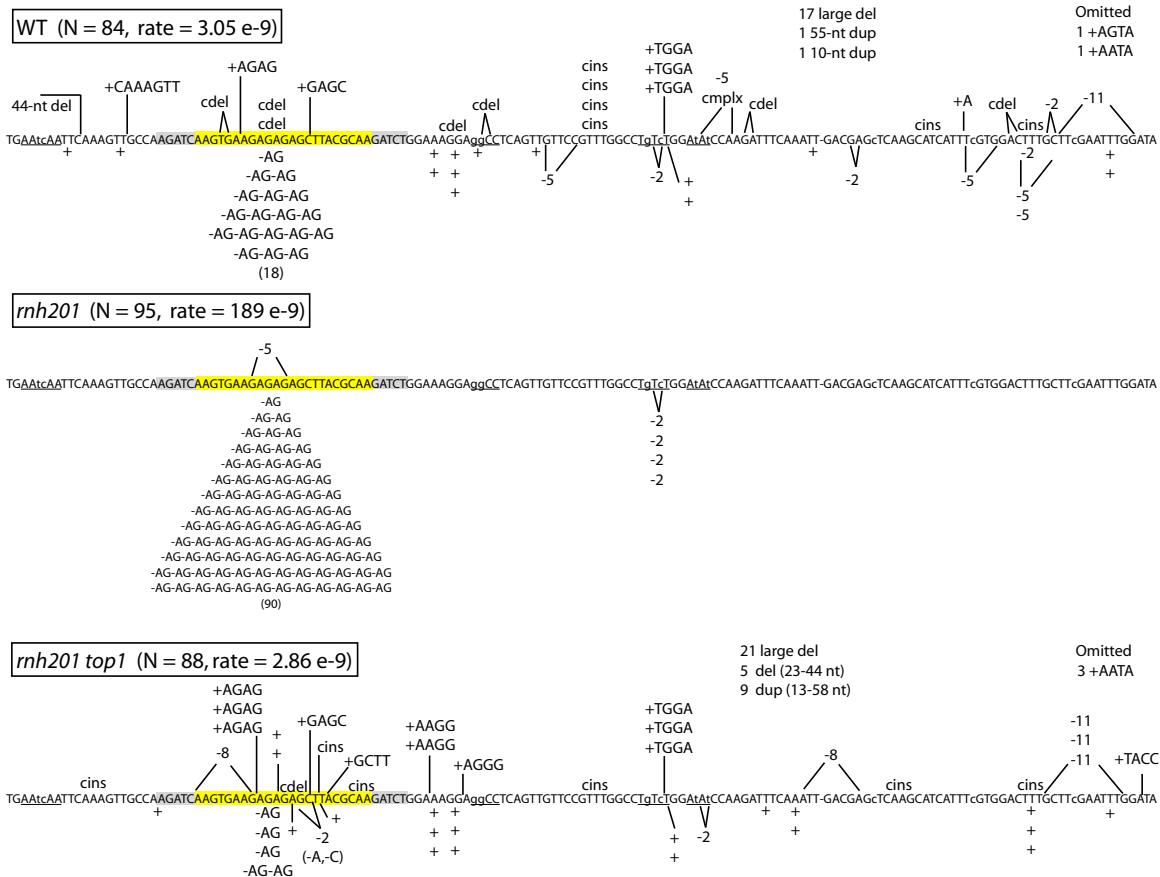
### A *rnh201* (N = 185, rate = 26e-8)

1 ATGACAAATT CAAAAGAAGA CGCCGACATA GAGGAGAACG ATATGTACAA TGAGCCGTC ACAACCTCT TTCAAGCAGT TGAAGCTCA CAAACACACC  
101 ACAGACGTGG GTCAATACCA TTGAAAGATG AGAAAAGTAA AGAATTGTAT CCATTGCGCT CTTCCCGAC GAGAGTAAT GGCGAGGATA CGTTCTCAT  
201 GGAGGATGGC ATAGGTGATG AAGATGAAG AGAAGTACAG AACGCTGAAG TGAAAGAGA GCTTAAGCAA AGACATATTG GTATGATTGC CCTTGGTGGT  
301 ACTATTGGTA CAGGTCTTTT CATTGGTTA TCCACACCTC TGACCAACGC CGGCCAGTG GGCCTCTTA TATCATATT ATTATGGGT TCTTGGCAT  
401 ATTCTGTAC GCAGTCCTTG GGTGAAATGG CTACATTCA CCCTGTTACA TCCTCTTCA CAGTTTCTC ACAAAAGATTC CTTCTCCAG CATTGGTGC  
501 GGCCAATGGT TACATGTATT GGTGTTCTTG GGCAATCACT TTTGCCCTGG AACTTAGTGT AGTGGCCAA GTCATTCAAT TTGGACGTA CAAAGTCCA  
601 CTGGCGGCAT GGATTAGTAT TTTTGGGTA ATTACACAA TAATGAACCT GTTCCCTGTC AAATATTACG GTGAATTGCA GTTCTGGTC GCTTCCATCA  
701 AAGTTTAGC CATTATCGG TTCTAATAT ACTGTTTTG TATGGTTGT GGTGCTGGGG TTACCGCCC ACTTGGATTG CGTTATTGGA GAAACCCAGG  
801 TGCCCTGGGT CCAGGTATAA TATCTAAGGA TAAAAACGAA GGGAGGTTCT TAGGTTGGGT TTCCCTTTG ATTAACGCTG CCTTCACATT TCAAGGTACT  
901 GAACTAGTTG GTATCACTGC TGTTGAAGCT GCAAACCCCA GAAAATCCGT TCCAAGAGCC ATCAAAAG TTGTTTCCG TATCTAACCC TTCTACATTG  
1001 GCTCTCTATT ATTCAATTGGA CTTTAGTTC CATAACAATGA CCCTAAACTA ACACAATCTA CTTCTACGT TTCTACTTCT CCCTTTATTA TTGCTATTGA  
1101 GAACTCTGGT ACAAAAGTTT TGCCACATAT CTTCAACGCT GTTATCTAA CAACCAATTAT TTCTGGCGCA AATTCAAATA TTACGTTGG TTCCCGTATT  
1201 TTATTTGGTC TATCAAAGAA CAAGTTGGCT CCTAAATTCC TGTCAGGAC CACCAAAGGT GGTGTTCCAT ACATTGAGT TTTCGTTACT GCTGCATTG  
1301 GCGCTTGGC TTACATGGAG ACATCTACTG GTGGTGACAA AGTTTCGAA TGGCTATTAA ATATCACTGG TGGTGCAGGC TTTTTGCA GTTATTAT  
1401 CTCATCTCG CACATCAGAT TTATGCAAGC TTGAAATAC CGTGGCATCT CTCGTGACGA GTTACCATTT AAAGCTAAAT TAATGCCGG CTTGGCTTAT  
1501 TATGCGGCCA CATTATGAC GATCATTATC ATTATTCAAG GTTTCACGGC TTTGCACCA AAATTCAATG GTGTTAGCCT TGCTGCCGCC TATATCTCTA  
1601 TTTCCCTGTT CTTAGCTGTT TGATCTTAT TTCAATGCAT ATTCAAGATGC AGATTATTT GGAAGATTGG AGATGTCGAC ATCGATTCCG ATAGAAGAGA  
1701 CATTGAGGCA ATTGTATGGG AAGATCATGA ACCAAAGACT TTTGGACA AATTGGAA TGTTGTAGCA TAG

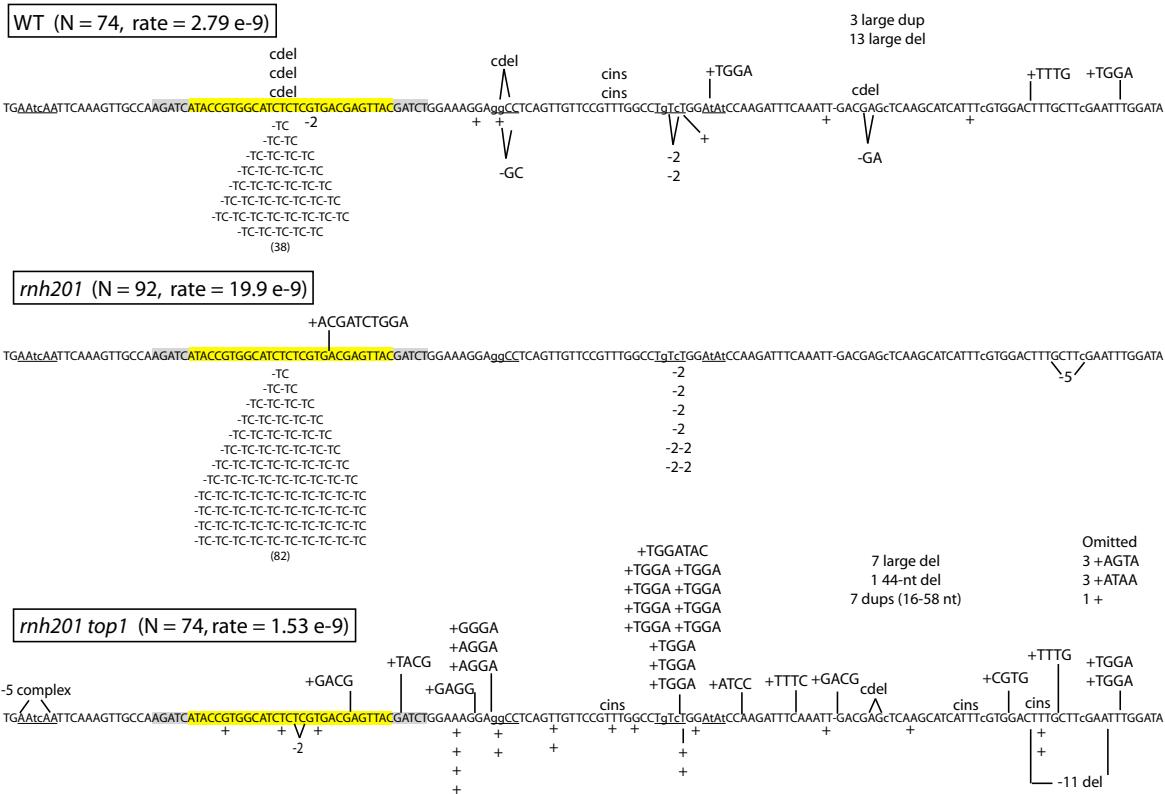
### B *rnh201 top1* (N = 165, rate = 18e-8)

1 ATGACAAATT CAAAAGAAGA CGCCGACATA GAGGAGAACG ATATGTACAA TGAGCCGTC ACAACCTCT TTCAAGCAGT TGAAGCTCA CAAACACACC  
101 ACAGACGTGG GTCAATACCA TTGAAAGATG AGAAAAGTAA AGAATTGTAT CCATTGCGCT CTTCCCGAC GAGAGTAAT GGCGAGGATA CGTTCTCAT  
201 GGAGGATGGC ATAGGTGATG AAGATGAAG AGAAGTACAG AACGCTGAAG TGAAAGAGA GCTTAAGCAA AGACATATTG GTATGATTGC CCTTGGTGGT  
301 ACTATTGGTA CAGGTCTTTT CATTGGTTA TCCACACCTC TGACCAACGC CGGCCAGTG GGCCTCTTA TATCATATT ATTATGGGT TCTTGGCAT  
401 ATTCTGTAC GCAGTCCTTG GGTGAAATGG CTACATTCA CCCTGTTACA TCCTCTTCA CAGTTTCTC ACAAAAGATTC CTTCTCCAG CATTGGTGC  
501 GGCCAATGGT TACATGTATT GGTGTTCTTG GGCAATCACT TTTGCCCTGG AACTTAGTGT AGTGGCCAA GTCATTCAAT TTGGACGTA CAAAGTCCA  
601 CTGGCGGCAT GGATTAGTAT TTTTGGGTA ATTACACAA TAATGAACCT GTTCCCTGTC AAATATTACG GTGAATTGCA GTTCTGGTC GCTTCCATCA  
701 AAGTTTAGC CATTATCGG TTCTAATAT ACTGTTTTG TATGGTTGT GGTGCTGGGG TTACCGCCC ACTTGGATTG CGTTATTGGA GAAACCCAGG  
801 TGCCCTGGGT CCAGGTATAA TATCTAAGGA TAAAAACGAA GGGAGGTTCT TAGGTTGGGT TTCCCTTTG ATTAACGCTG CCTTCACATT TCAAGGTACT  
901 GAACTAGTTG GTATCACTGC TGTTGAAGCT GCAAACCCCA GAAAATCCGT TCCAAGAGCC ATCAAAAG TTGTTTCCG TATCTAACCC TTCTACATTG  
1001 GCTCTCTATT ATTCAATTGGA CTTTAGTTC CATAACAATGA CCCTAAACTA ACACAATCTA CTTCTACGT TTCTACTTCT CCCTTTATTA TTGCTATTGA  
1101 GAACTCTGGT ACAAAAGTTT TGCCACATAT CTTCAACGCT GTTATCTAA CAACCAATTAT TTCTGGCGCA AATTCAAATA TTACGTTGG TTCCCGTATT  
1201 TTATTTGGTC TATCAAAGAA CAAGTTGGCT CCTAAATTCC TGTCAGGAC CACCAAAGGT GGTGTTCCAT ACATTGAGT TTTCGTTACT GCTGCATTG  
1301 GCGCTTGGC TTACATGGAG ACATCTACTG GTGGTGACAA AGTTTCGAA TGGCTATTAA ATATCACTGG TGGTGCAGGC TTTTTGCA GTTATTAT  
1401 CTCATCTCG CACATCAGAT TTATGCAAGC TTGAAATAC CGTGGCATCT CTCGTGACGA GTTACCATTT AAAGCTAAAT TAATGCCGG CTTGGCTTAT  
1501 TATGCGGCCA CATTATGAC GATCATTATC ATTATTCAAG GTTTCACGGC TTTGCACCA AAATTCAATG GTGTTAGCCT TGCTGCCGCC TATATCTCTA  
1601 TTTCCCTGTT CTTAGCTGTT TGATCTTAT TTCAATGCAT ATTCAAGATGC AGATTATTT GGAAGATTGG AGATGTCGAC ATCGATTCCG ATAGAAGAGA  
1701 CATTGAGGCA ATTGTATGGG AAGATCATGA ACCAAAGACT TTTGGACA AATTGGAA TGTTGTAGCA TAG

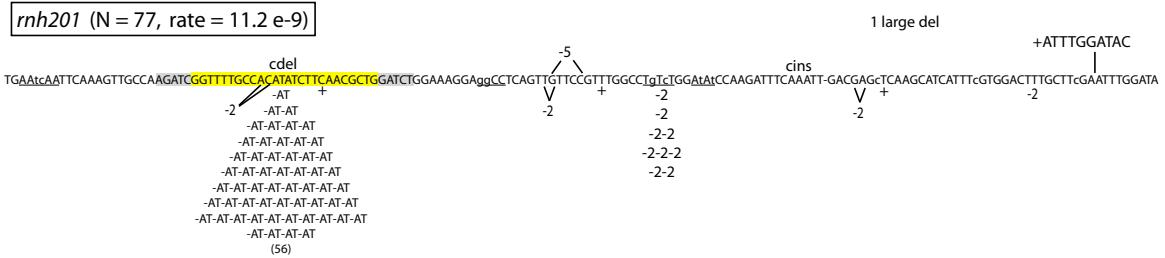
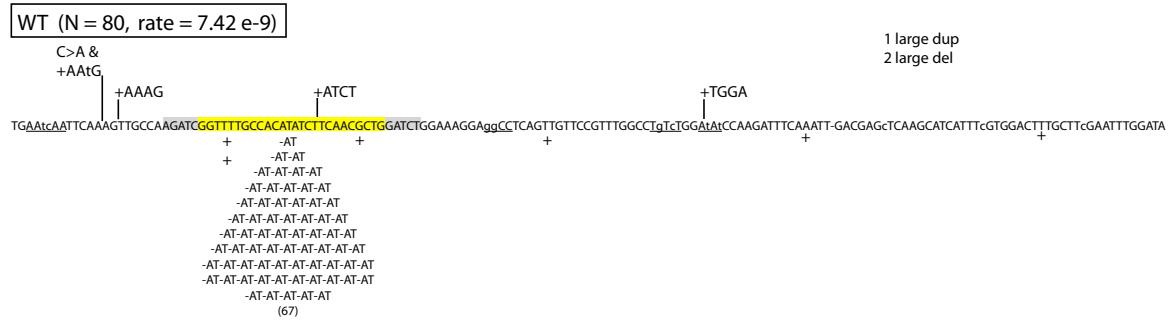
## A *lys2ΔA746NR,(AG)4* reversion



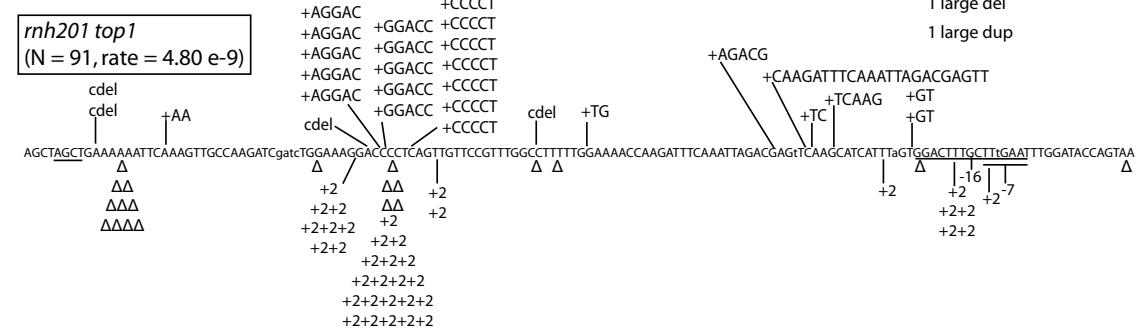
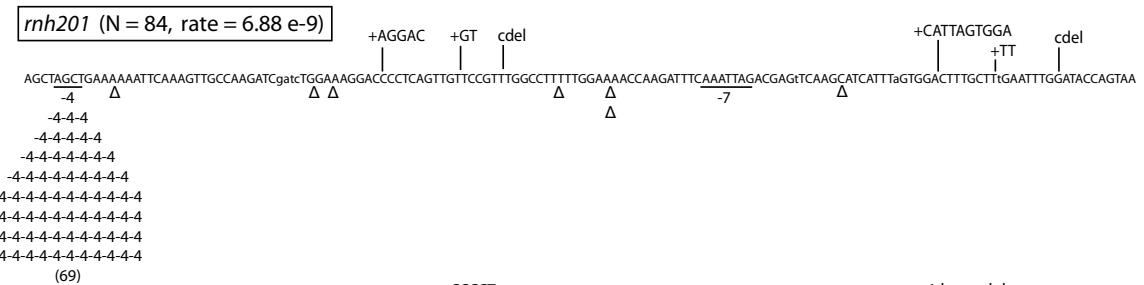
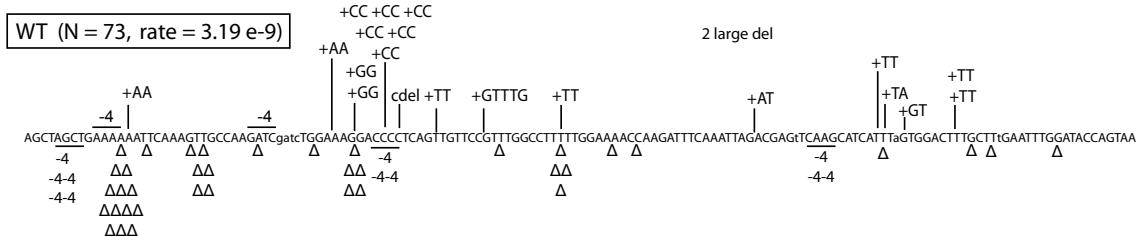
## B *lys2ΔA746NR,(TC)3* reversion



## C *lys2ΔA746NR,(AT)2* reversion



## D *lys2ΔBgl* reversion



**A**

$(AG)_4$	<p>NTS 5' -AACGCTGAAGTGA AGAGAGAG CTTAAGCAA            TS            TTGCGACTTCACT TCTCTCTC GAATTCGTT-5'</p> <p style="text-align: center;">a      b ↑      ↓ c      d</p>
$(AT)_2$	<p>NTS 5' -ACAAAGGTTTGCCAC ATAT CTTCAACGCT            TS            TGTTTCCAAAACGGTG TA TA GAAGTTGCGA-5'  <p style="text-align: center;">e      f      g</p> </p>
$(TC)_3$	<p>NTS 5' -TTTGAATACCGTGGCA TCTCTC GTGACGAGT            TS            AAACTTATGGCACCGT AGAGAG CACTGCTCA-5'  <p style="text-align: center;">h</p> </p>

**B**